CLAIMS

We claim:

1. An isolated protein, comprising an amino acid sequence having at least 80% sequence identity to SEQ ID NO: 2, wherein the protein has taxoid oxygenase activity.

2. The isolated protein of claim 1, comprising an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2.

3. The isolated protein of claim 1, comprising an amino acid sequence as set forth in SEQ ID NO: 2.

4. The isolated protein of any one of claims 1-3, wherein the taxoid oxygenase activity comprises taxadiene hydroxylation.

5. An isolated nucleic acid molecule encoding the protein of claim 1, 2, 3, or 4.

6. An isolated nucleic acid of claim A5 encoding the protein of claim 1.

7. An isolated nucleic acid molecule according to claim 5 comprising a sequence as set forth in SEQ ID NO: 1.

8. A recombinant nucleic acid molecule, comprising a promoter sequence operably linked to a nucleic acid molecule according to claim 5.

9. A cell transformed with a recombinant nucleic acid molecule according to claim 8.

10. The cell of claim 9, wherein the cell is a plant cell, an insect cell, a bacterium, or a yeast cell.

11. The cell of claim 10, wherein the cell is a plant cell from the genus *Taxus*.

12. The cell of claim 9, wherein the cell is an isolated cell!

13. A non-human transgenic organism, comprising a recombinant nucleic acid molecule according to claim 8.

14. The organism of claim 13, wherein the organism is a plant.

15. The organism of claim 14, wherein the plant is from the genus *Taxus*.

16. An isolated nucleic acid molecule that hybridizes under high stringency conditions with a nucleic acid probe comprising at least 600 base pairs of SEQ ID NO: 1, wherein the nucleic acid molecule encodes a protein having taxoid oxygenase activity.

17. An isolated nucleic acid molecule according to claim 16 that hybridizes under very high stringency conditions.

18. The protein encoded by the nucleic acid molecule of claim 17.

19. A recombinant nucleic acid molecule, comprising a promoter sequence operably linked to the nucleic acid molecule of claim 16.

20. A cell transformed with a recombinant nucleic acid molecule according to claim 19.

21. The cell of claim 20, wherein the cell is a plant cell, an insect cell, a bacterium, or a yeast cell.

22. The cell of claim 21, wherein the cell is a plant cell from the genus *Taxus*.

23. The cell of claim 21, wherein the cell is an isolated cell.

24. A non-human transgenic organism, comprising a recombinant nucleic acid molecule according to claim 19.

25. The organism of claim 24, wherein the organism is a plant.

26. The organism of claim 25, wherein the plant is from the genus *Taxus*.

27. An isolated nucleic acid molecule, comprising a sequence having at least 80% sequence identity with SEQ ID NO: 1, wherein the nucleic acid molecule encodes a protein having taxoid oxygenase activity.

28. An isolated nucleic acid molecule according to claim 27, comprising the sequence set forth in SEQ ID NO: 1.

29. The protein encoded by the nucleic acid molecule of claim 28.

30. A recombinant nucleic acid molecule, comprising a promoter sequence operably linked to the nucleic acid molecule of claim 27.

31. A cell transformed with a recombinant nucleic acid molecule according to claim 30.

32. The cell of claim 31, wherein the cell is a plant cell, an insect cell, a bacterium, or a yeast cell.

33. The cell of claim 32, wherein the cell is a plant cell from the genus *Taxus*.

34. The cell of claim 32, wherein the cell is an isolated cell.

35. A non-human transgenic organism, comprising a recombinant nucleic acid molecule according to claim 30.

36. The organism of claim 35, wherein the organism is a plant.

37. The organism of claim 36, wherein the plant is from the genus *Taxus*.

38. A method of identifying a nucleic acid sequence that encodes a taxoid oxygenase, comprising:

(a) hybridizing a probe to a nucleic acid sequence under high stringency conditions, wherein the probe comprises at least 600 contiguous nucleotides of SEQ ID NO: 1; and

(b) determining that a protein encoded by the nucleic acid sequence is capable of oxidizing a taxoid substrate, wherein a protein capable of oxidizing a taxoid substrate is identified as a taxoid oxygenase.

39. The method of claim 38, wherein step (a) is performed under very high stringency conditions.

40. The method of claim 38, wherein oxidizing the taxoid substrate comprises hydroxylating the taxoid substrate.

41. A method of hydroxylating a substrate, comprising:

contacting a substrate with at least one oxygenase comprising an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2; and

allowing the oxygenase to oxidize the substrate.

42. The method of claim 41, wherein oxidation of the substrate comprises hydroxylation of the substrate.

43. The method of claim 41, wherein the substrate is a taxoid.

44. The method of claim 43, wherein the taxoid is paclitaxel or a paclitaxel intermediate.

45. The method of claim 43, wherein the taxoid is a taxadiene.

46. The method of claim 45, wherein the taxadiene is taxa-4(5),11(12)-diene or taxa-4(20),11(12)-diene.

47. The method of claim 43, 44, 45, or 46, wherein the oxygenase hydroxylates position C5 of the taxoid.

48. The method of claim 41, wherein the amino acid sequence is SEQ ID NO: 2.

49. The method of claim 41, 43, 44, 45, 46, or 48, wherein the oxygenase is expressed in an isolated cell or in a transgenic plant, bacterium, insect, fungus or yeast, and the hydroxylation of the substrate occurs *in vivo*.

50. The method of claim 49, wherein the substrate is an exogenous substrate, which is fed to the isolated cell, transgenic plant, transgenic bacterium, transgenic insect, transgenic fungus or transgenic yeast.

51. A method for increasing paclitaxel yield in a cell, comprising introducing at least one nucleic acid molecule of claim 6, 16, or 27 into a paclitaxel-producing cell, wherein an amount of paclitaxel produced by the cell increases following introduction of the at least one nucleic acid molecule into the cell.

52. The method of claim 51, wherein introducing the at least one nucleic acid molecule into the cell comprises transfecting the cell with the at least one nucleic acid molecule.

53. The method of claim 51, wherein the at least one nucleic acid molecule comprises a sequence as set forth in the protein-coding region of SEQ ID NO: 1.

54. The method of claim 51, wherein the paclitaxel produced by the cell increases by at least two fold.

55. The method of claim 51, wherein the cell is a *Taxus* cell.

56. The method of claim 51, further comprising introducing into the cell an additional nucleic acid molecule selected from the group consisting of:

(a) a first nucleic acid sequence encoding a protein having taxadiene synthase activity;

(b) a second nucleic acid sequence encoding a protein having taxadien-5-ol transacylase activity;

(c) a third nucleic acid molecule encoding a protein having taxadien-2-ol transacylase activity;

(d) one or more fourth nucleic acid molecules encoding one or more proteins having taxoid oxygenase activity;

(e) a fifth nucleic acid molecule encoding a protein having 10-deacetylbaccatin III-10-O-acetyl transferase activity;

(f) a sixth nucleic acid molecule encoding a protein having taxoid 13-phenylpropanoyltransferase activity;

(g) a seventh nucleic acid molecule encoding a protein having 3'-N-debenzoyltaxol N-benzoyltransferase activity; and

(h) combinations of (a), (b), (c), (d), (e), (f), or (g).

57. The method of claim 56, wherein:

(a) the first nucleic acid molecule comprises a nucleic acid sequence having at least 90% sequence identity to the protein-coding region of SEQ ID NO: 19 and encodes a protein having taxadiene synthase activity;

(b) the second nucleic acid molecule comprises a nucleic acid sequence having at least 90% sequence identity to the protein-coding region of SEQ ID NO: 21 and encodes a protein having taxadien-5-ol transacylase activity;

(c) the third nucleic acid molecule comprises a nucleic acid sequence having 90% sequence identity to the protein-coding region of SEQ ID NO: 23 and encodes a protein having taxadien-2-ol transacylase activity;

(d) the one or more fourth nucleic acid molecules comprise a nucleic acid sequence having 90% sequence identity to any one of the sequences set forth in SEQ ID NOs: 3, 5, 7, 9, 11, 13, 15, 17, and 40 and encode a protein having taxoid oxygenase activity;

(e) the fifth nucleic acid molecule comprises a nucleic acid sequence having 90% sequence identity to the protein-coding region of SEQ ID NO: 34 and encodes a protein having 10-deacetylbaccatin III-10-O-acetyl transferase activity;

(f) the sixth nucleic acid molecule comprises a nucleic acid sequence having 90% sequence identity to the protein-coding region of SEQ ID NO: 36 and encodes a protein having taxoid 13-phenylpropanoyltransferase activity; and

(g) the seventh nucleic acid molecule comprises a nucleic acid sequence having 90% sequence identity to the protein-coding region of SEQ ID NO: 38 and encodes a protein having 3'-N-debenzoyltaxol N-benzoyltransferase activity.

58. The method of claim 56, wherein the one or more fourth nucleic acid molecules encode one or more proteins having taxoid7β-hydroxylase activity, taxoid 14β-hydroxylase activity, taxoid 10β-hydroxylase activity or taxoid 13α-hydroxylase activity.

59. The method of claim 56, wherein introducing the additional nucleic acid molecule into the cell comprises transfecting the cell with the additional nucleic acid molecule.

60. The method of claim 56, wherein the additional nucleic acid molecule is selected from the group consisting of:

(a) a first nucleic acid molecule comprising a nucleic acid sequence having at least 90% sequence identity to the protein-coding region of SEQ ID NO: 19, wherein the first nucleic acid sequence encodes a protein having taxadiene synthase activity;

(b) a second nucleic acid molecule comprising a nucleic acid sequence having at least 90% sequence identity to the protein-coding region of SEQ ID NO: 21, wherein the second nucleic acid sequence encodes a protein having taxadien-5-ol transacylase activity;

(c) a third nucleic acid molecule comprising a nucleic acid sequence having 90% sequence identity to the protein-coding region of SEQ ID NO: 23, wherein the third nucleic acid sequence encodes a protein having taxadien-2-ol transacylase activity; and

(d) combinations of (a), (b), or (c).

61. The method of claim 60, wherein the additional nucleic acid molecule is selected from the group consisting of:

(a) a first nucleic acid molecule comprising the protein-coding region of SEQ ID NO: 19;

(b) a second nucleic acid molecule comprising the protein-coding region of SEQ ID NO: 21;

(c) a third nucleic acid molecule comprising the protein-coding region of SEQ ID NO: 23; and

(d) combinations of (a), (b), or (c).

62. The method of claim 51, wherein the cell is a cell line.

63. An antibody or antibody fragment that binds the protein of claim 1, 18, or 29.

64. The antibody of claim 63, wherein the antibody is a monoclonal antibody.

65. The antibody fragment of claim 63, wherein the antibody fragment is a Fab, F(ab)2, or Fv fragment, or a combination thereof.